

1/27

FIG. 1

NUCLEIC ACID AND PREDICTED PROTEIN SEQUENCE OF DZ2

1 GGCACGAGCAGAATCGAAGATGGCAACAATAATCCATGGGAGATATCGAGAAATAAGAA 60
M A T K S M G D I E K I K K

61 GAACTAAACGTGTTGATCGTCGATGATCCACTAACCTTATAATTATGAGAAGAT 120
K L N V L I V D D D P L N L I I H E K I

121 CATCAAAGCGATTGGGGGTATTTACACAGACAGCGGAATAACGGTGAGGAGGCAGTAATCAT 180
I K A I G G I S Q T A N N G E E A V I I
↓ DZ2FL→

181 CCACCGTGACGGCGGCTCATCTTTTGACCTTATCCTAATGGATAAAGAAATGCCCGAGAG 240
H R D G G S S F D L I L M D K E M P E R

241 GGATGGTGTTCGACAACCTAAGAAGCTAAGAGAAATGGAAGTGAATGATTGTTGG 300
D G V S T T K K L R E M E V K S M I V G

301 GGTGACTTCACTGGCTGACAATGAAGAGGAGCGCAGGGCTTTTCATGGAAGCTGGACTTAA 360
V T S L A D N E E E R R A F M E A G L N

2/27

```

361 CCATTGCTGGCAAAACCGTTAACCAAGACAAGATCATCCCTCTCATTAACCAACTCAT 420
    H C L A K P L T K D K I I P L I N Q L M
    ← DZ2RL
421 GGATGCTTGATGGATATATATTTATATGGAACACACATAATAACGTCTAAGTGTG 480
    D A *
481 TATGATGCATAGATAGATACTTGCATGTTGTGTTTAGAATTTAGGGTTCTTTATCGTCCGT 540
    HindIII
541 GATATATAATCATGTAAGTTGTTGCTTTAAGCTTATAAAATATTAAATAAGGGTTTCCT 600
601 CTACC
    
```

The primer sites for DZ2FL and DZ2RL are underlined, as are the recognition sequences for *NcoI* and *HindIII* restriction enzymes. Shown in bold are the conserved amino acid residues required for phosphorylation. The extent of the original PCR product isolated by differential display is shown by ↓.

FIG. 1 CONT'D

3/27

Alignment of the predicted protein sequence of DZ2 with those of
bacterial response regulator proteins.

	1				50
DZ2	MATKSMGDIE	KIKKKLNVLI	VDDDPLNLII	HEKIIKAIG.	GISQTANNGE
OMPRMQENYKILV	VDDDMRLRAL	LERYLTEQGF	.QVRSVANAE
PHOBMARRILV	VEDEAPIREM	VCFFVLEQNGF	.QPVEAEDYD
NTRCMQRGIVWV	VDDSSIRWV	LERALAGAGL	.TCTTFENG
SPOOFMMNEKILI	VDDQYGIRIL	LNEVFNKEGY	.QTFTAANG
CHEY	MADKELKFLV	VDDFSTMRRI	VRNLLKELGF	NNVEEAEDGV
ETRLKVLV	MDENGVSVMV	TKGLLVHLGC	EVTTVSSNEE
	51				100
DZ2	EAVIIHRDGG	SSFDLILMDK	EMPERDGVST	TKKLREMEVK	SM..IVGVTS
OMPR	QMDRLLTR..	ESFHLMVLDL	MLPGEDGLSI	CRRLRSQS..	NPMPIIMVTA
PHOB	SAVNQLNE..	PWPDILLDW	MLPGSGIQF	IKHLKRESMT	RDIPVVMLTA
NTRC	EVLAALAS..	KTPDVLLSDI	RMPGMDGLAL	LKQIKQ..RH	PMLPVIIMTA
SPOOF	QALDIVTK..	ERPDLVLLDM	KIPGMDGIEI	LKRMKV..ID	ENIRVIIMTA
CHEY	DALNKLQA..	GGYGFVISDW	NMPNMDGLEL	LKTIRADGAM	SALPVLIMVTA
ETRCLRVS	HEHKVVFMDV	CMPGVENYQI	ALRI.....	.HXPLLVALS
	101				150
DZ2	LADNEEERRA	FMEAGLNHCL	AKPLTKDKII	PLINQLMDA.
OMPR	KGEEVDRIVG	.LEIGADDYI	PKPFNPPELL	ARIRAVLRRQ	ANELPGAPS.
PHOB	RGEEDRVRG	.LETGADDYI	TKPFSPKELV	ARIKAVMRRI	SPM.....
NTRC	HSDLDAVSA	.YQQGAFDYL	PKPFDIDEAV	ALVERAISHY	QEQQOPRNI
SPOOF	YGELDMIQES	.KELGALTFH	AKPFDIDEIR	DAVKKYLPLK	SN.....
CHEY	EAKKENIIAA	.AQAGASGYV	VKPFPTATLE	EKLNKIFEKL	GM.....
ETR	GNTDKSTKEK	CMSFGLDGLV	LKPVS LDNIR	DVLSDLL...
	151				
DZ2				
OMPR	OmpR = <i>E. coli</i>	(Involved in osmoregulation)		
PHOB	PhoB = <i>E. coli</i>	(Involved in phosphate utilisation)		
NTRC	VNGP	NtrC = <i>S. typhimurium</i>	(Involved in nitrogen utilisation)		
SPOOF	SpoOF = <i>B. subtilis</i>	(Involved in sporulation)		
CHEY	CheY = <i>E. coli</i>	(Involved in chemotaxis)		
ETR	ETR = <i>A. thaliana ETR1</i> gene	encoding an ethylene receptor (partial amino acid sequence)		

The predicted protein sequence of DZ2 is shown in bold as are the conserved amino acid residues required for phosphorylation of the protein

FIG. 2

4/27

Expression analysis of DZ2 in various plant organs using Northern

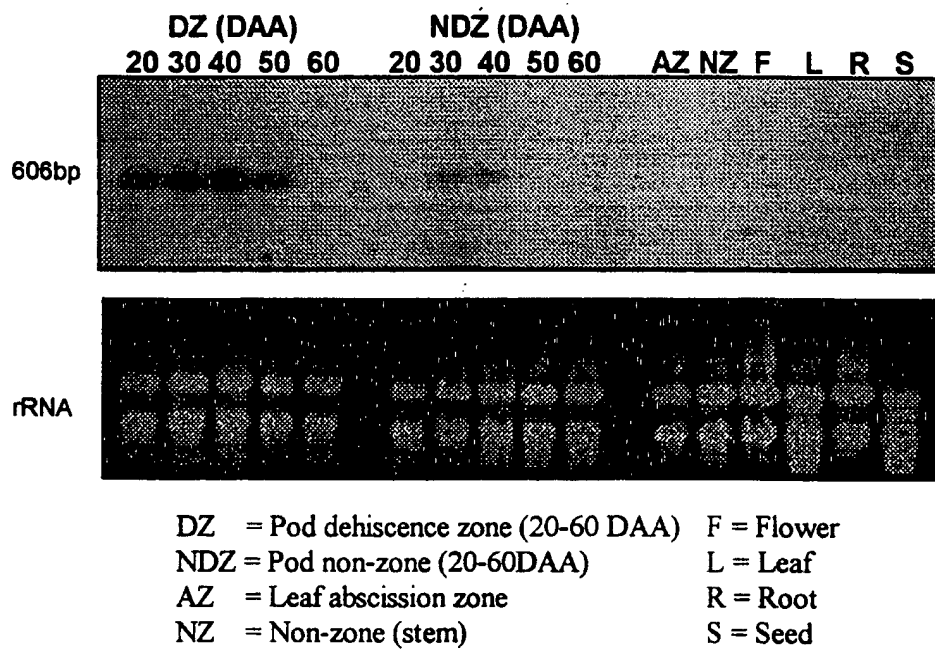
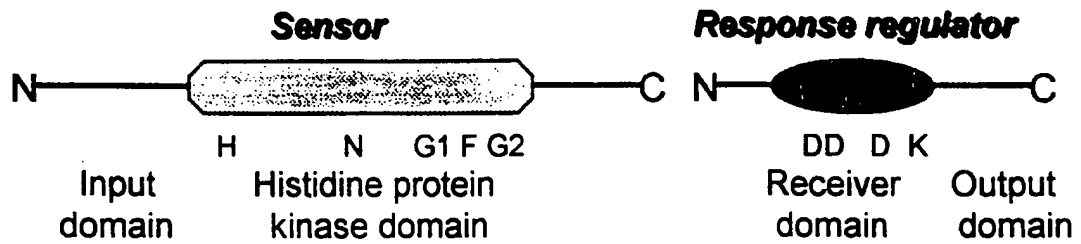


FIG. 3

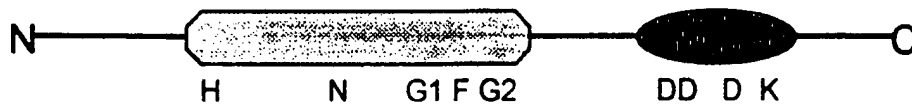
5/27

Comparison of bacterial two-component regulatory systems with DZ2

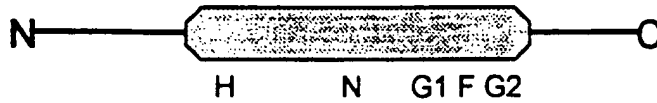
Bacterial two-component system



***Arabidopsis thaliana* ETR1 gene**



***Arabidopsis thaliana* ERS gene**



DZ2



FIG. 4

6/27

FIG. 5

TATATAAATACGGTTTAAACAGATATGTTCTGGTTATATAAATGTAATTCNATGTGCCNNTCAANTTTTATTTTNATNGT 78

TNCTACTAGGGACATTAGTTTTTAAACNTTTTATATATCATGTGAACAAAAACNTTTTATATNCAACTATGA 156

GCAATTATCTTATAGTGTCTTTTCCAGAAATTTGACGACAACCTAACTAAACAATTTAATTTGACGTTAGTT 234

AAGTAAATTTATATAGATGATAAATTTGAGCAAGCACATTACGAACTGCGGATCAAGGAGAGTCACAATTTAATCTTA 312

GGCTCTAGACGAACCTGCGGAGCAAGG -> DZ2BGENF

XbaI

CGTTATACACAAAAATTATCTAAATACTATATATATATACAGCTGCAATGCTACGATAATGATCAAAATGTTTATGTACTT 390

TTCAGCGAAAAATCTTTGTGCGCCATACATTACTGTGTTAATGAATCATTTAAATATGTGAAGGAGGAAAAAGAGTACAAAA 468

GGAGTTTGTGAGGCAATTCGCAGACACTGAAATGTGAATAATAATAAGGAATTGCCGAATTGATTTCTAGTTGGT 546

GAAGTGGTGAAAAATTGTATGTCCATTGCTTATAAACTATAAAATATAATATNTTNATATTATCACTNTGGACATTAG 624

TNNGATAGACCCCTAGCTAAAAATTTTAAAAAATTATACATTTCATTTTCTNAAGTACCAAACTTAATTTATCACAATCGGA 702

TAAAAATTGTTTAAGAAACCAATTACAAACTCAGCTTGTGGACTCTGAGAGAAACTAAGAGCTAGACATACGGTTAGTAG 780

TGTAGCCGCATTTTATATGCTTAATTTGCTTAAGCATGACTTCTATGCTCCTTGATGATATTTTATTAATATCCTAG 858

GACATATGGATTTGATAAAGATCTTATCAACCTTTCACAAGACCAATTAGCTCAACAACAAAACTACTGAAAGTATAT 936

7/27

AAATCTTGGTTACAGAAATTCCTTATGCCAAAATATCATATAATATAGAAATTCGGTTATGATTAAAGATCAATTATT	1013
AATTAATATAATTTTTTCACCTTTTGTCTTATGTATCTTATAGTATTTGTTTACCATATTTGACCGAATGGTGTGCATATTA	1092
GTTTGGTAAGACAACACTCAGTTGCAACGATGCAGATTACATTTTCAGGAAGATTCATGTAAGAAAGATAATTTTCGCTTTTGT	1170
GGTGTGAAAAAATATGCCCTCTTTTCACCTTTTTTTTCAACTATAAAATTTTCGATCGATGTATCTACGTTCTTAAACACAATTCAC	1248
AAATCTTCTTTAGAAATCCAAAATTTGTAAGCCGCTTTCTAAATCTCTTTCTCAGTATACATATATGTAATAATGTATGCATATA	1326
TTATTATTTCATAATACAAACACGAACCCATGCATGCAAGAGATAGTTACACGGCTCATAAACAACACAAAAAACAATA	1404
CGCATGCATTAGAACACCTTGTATGTTAATTTTCCATAATGTTTTTTGCATAAAACATTCCTCGTTTTTAAATTAGCTTCTTTTTT	1482

NcoI

<- GTTCGAAGAAAAAACCGACCATGGCAG DZ2BGENR AACCAAG
GTGTGAAGATTGTTCGAAGAAAAAACCGAAGATGGCAACGTCACATCCACGGAGATATCAAGAAACCAAG 1560
M A T T S T S T G D I K K T K

TCAGTAGAAGTG -> DZ2BFL
TCAGTAGAAGTGAAGAAGAACTTAACGTGTTGATCGTTGATGATACAGTAATTCGTAAACTTCACGAGAATATC 1638
SSV E V K K K L N V L I V D D D T V I R K L H E N I
ATCAAATCGATCGTGGAAATTCACAGACGGCTAAGAACGGTGAGGAGGCAGTGAACATCCACGCGACGGCAATGCA 1716
IKS I G G I S Q T A K N G E E A V N I H R D G N A ->

FIG. 5 CONT'D

8 / 27

FIG. 6

A) D22B cDNA sequence (top) -
D22 sequence (bottom)

TCGTCTNAT

Dz2 sequence (bottom)

GGCACGACGACGAATCGAAGATGGCAACAAATCCATGGGAGATATCGAGAAAATAAAGAAGAACTAAACGTGTTGATCGTC

TCGTCNAT

GATGATCCTGTAAACGGTAAACTTCACGAGATTATCATCAAAATCAATCGGTGA---ATTTCACAGACAGCTAAGAACGGTG
.....
GATGATGATCCACTAAACCTTTATAATTGATGAGAAGATCATCAAAAGCGATTGGGGGTATTTCACAGACAGCGAATAACGGTG

10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160

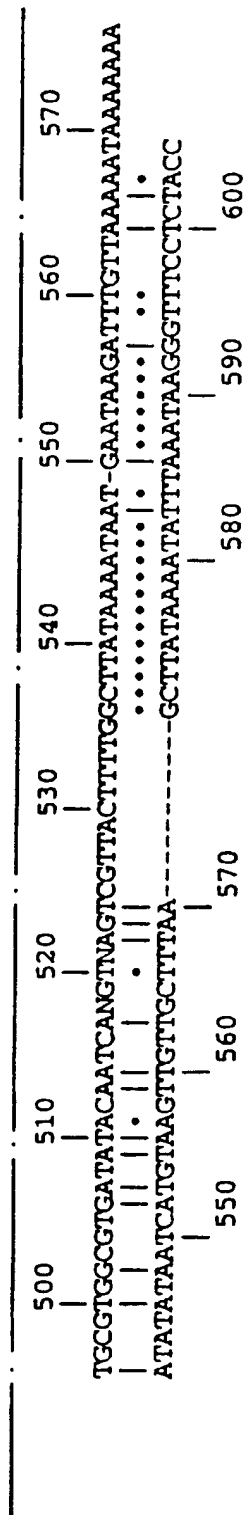
[illegible]

170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 |
 ACTTTCGCAACTAAGAGAAATGAAAGTGCATGCTATGATTATGGGGTGACGACACTGGCTGACCAATGAAGAG
 | | | | | | | | |
 TGTTCGCAACTAAGAGAAATGGAAGTGAAGTCAATGATTGTTGGGGTGACTTCACTGGCTGACCAATGAAGAG
 | 250 | 260 | 270 | 280 | 290 | 300 | 310 | 320 |

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIG. 6 CONT'D

10/27



**B) D22B peptide sequence (top) -
D22 peptide sequence (bottom)**

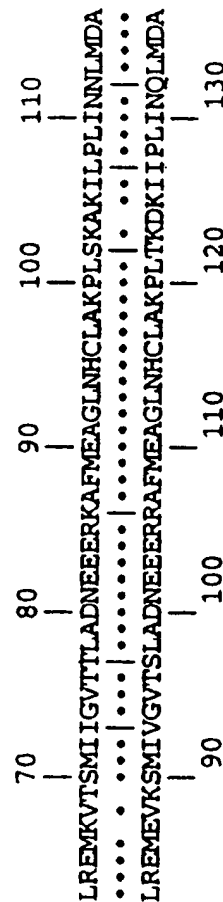
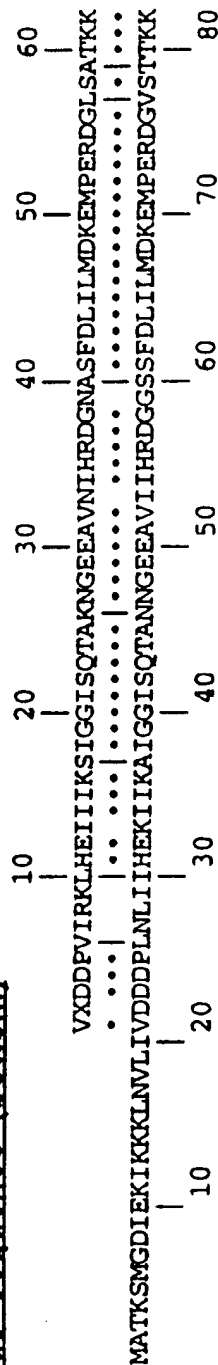
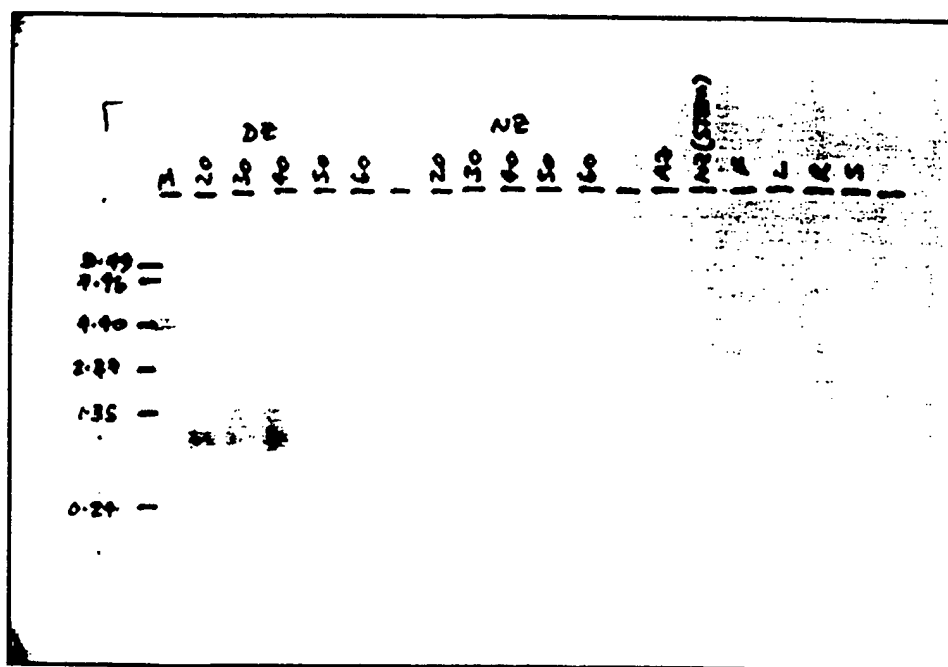


FIG. 6 CONT'D

11/27

FIG. 7

NORTHERN ANALYSIS OF EXPRESSION OF
DZ2B IN PODS AND OTHER TISSUES



DZ = POD DEHISCENCE ZONE (20 - 06 daa)
 NZ = POD NON-ZONE
 AZ = ABSCISSION ZONE
 F = FLOWER
 L = LEAF
 R = ROOT
 S = SEED

12/27

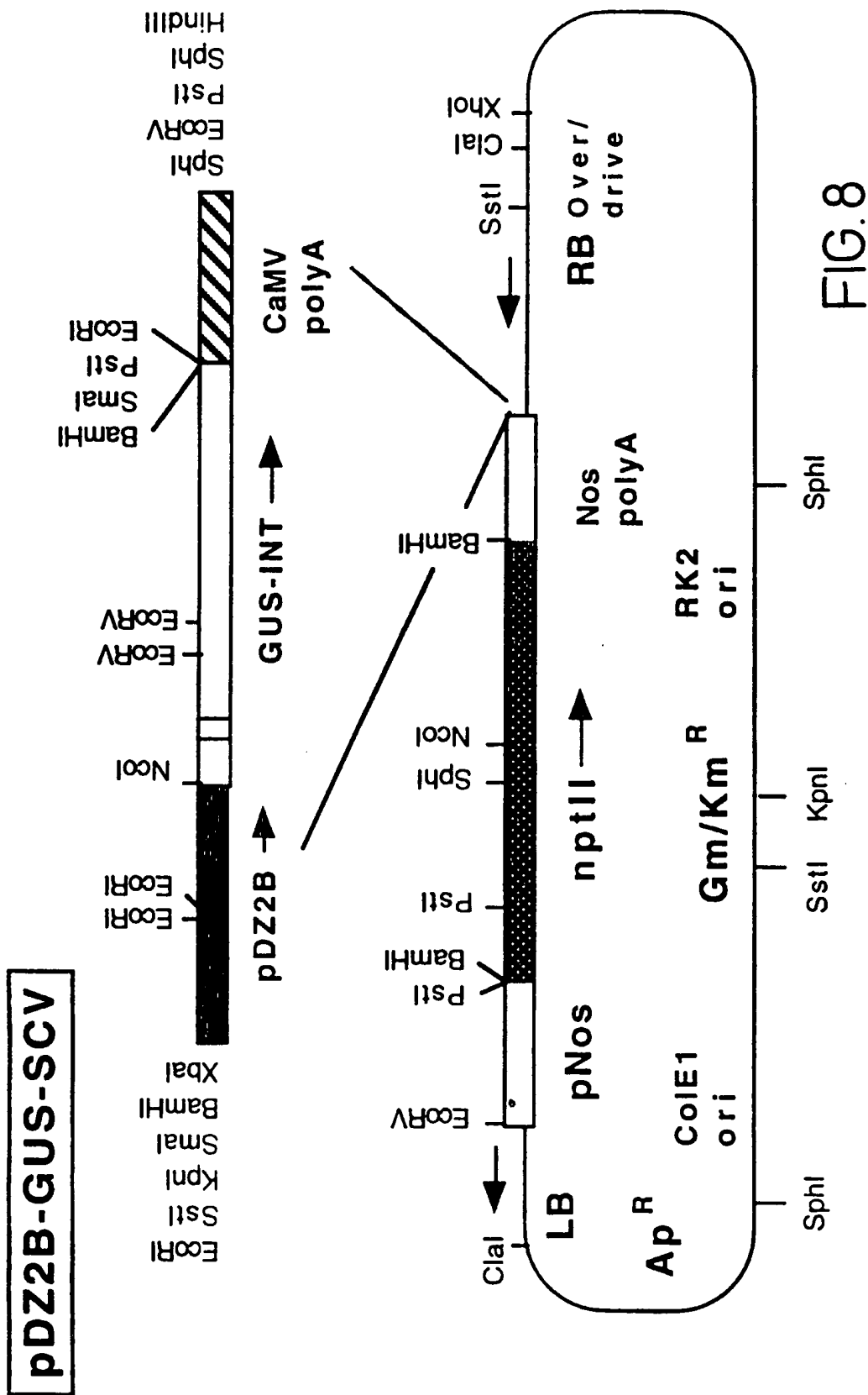


FIG.8

FIG. 9

ATATATGTGATACAGATACATCTATATACAAAATTAAACACGAAACCATACATGACGGTGTGATCACACACGCACACA 78

CATAGAAACATAAACACGCAATAATTCTATACAGTTTAATTTCATTTTTTAACTTACTTCTTTTTTTTGGTGAAGAT 156

TCTTGAGAGAAAGAAATCGAAGATGCGCAACAAATCCACCGAGGTACCGAGAAACCAAGTCGATAGAAGTGAAGA 234
M A T K S T G G T E K T K S I E V K K

AGAAACTAATCAACGTGTTGATCGTCGATGATCCATTAAACCGTAGACTCCACGAGATGATCATCAAAACGATCG 312
K L I N V L I V D D D P L N R R L H E M I I K T I G
<- CGATCG

<- CGAATAACGGTGAGGAGGCATTAATCA GW1
GAGGAATTTCTCAGACTGCAA AT3GW2
GAGGAATTTCTCAGACTGCAAAGAATGGCGAAGAGGCNGTGATCCTCCACCGTGACGGCGAAGCATCTTTTCGACCTTA 390
G I S Q T A K N G E E X V I L H R D G E A S F D L I

TTCTAATGGATAAGGAAATGCCCTGAGAGGGATGGAGTTTCGACAAATTAAGANGCTAAGAGAAATGAAAGGACGTCAA	468
L M D K E M P E R D G V S T I K X L R E M K G T S M	
TGATCGTTGGGTAACGTACGTAGCTGACCAAGAAGAGCGTAAGGCTTTTATGGAAGCTGGGCTCAACCATTTGCT	546
I V G V T S V A D Q E E E R K A F M E A G L N H C L	
TGAAAAACCCTTAACCAAGGCCAAGATCTTCCCGCTCATTAGCCACCTCTTCGATGCTTGATGATGAAGGCTCATT	624
E K P L T K A K I F P L I S H L F D A .	
AATGTATCTATATTTTCAATCATGAATCACCTACACGTGTATTGTGACACAAAAATCTGCATTGTGTGATATAGGG	702
TTTCTCATATCTATGTTGATTTATTTTCTTATCGTCCGAGGTAAATCATGCAAGTCATTTCTTTTGGCTAATAAAA	780
TATTAAAAAAGGTTTTCTCAAAAAAATAAAAAA	818

FIG. 9 CONT'D

15/27

FIG. 10

A) DZ2AT3 peptide sequence (top)
DZ2 peptide sequence (bottom)

10	20	30	40	50	60	70	80
MATKSTGGTEKTSIEVKKKLINVLIVDDDPNRRRLHEMI	IKTIGGISQTAKNGEEXVILHRDGEASFDLILMDKEMPERDG						
..
MATKSMGDIKIKKKLNVLIIVDDDPNLIIHEKIIKAIGGISQTANNGEAAV	IHRDGGSSFDLILMDKEMPERDG						
10	20	30	40	50	60	70	

90	100	110	120	130	140
VSTIKXLREMGTS	MIVGTVTSVADQEEERKAFMEAGLNHCL	LEKPLTKAKIFPLISHLFDA			
...
VSTTKKLREMEVK	SMIVGTVTSLADNNEERRAFMEAGLNHCL	AKPLTKDKIIP	LINQLMDA		
80	90	100	110	120	130

B) DZ2AT3 peptide sequence (top)
DZ2B peptide sequence (bottom)

10	20	30	40	50	60	70	80
MATKSTGGTEKTSIEVKKKLINVLIVDDDDPLNRRRLHEMI	IKTIGGISQAKNGEEXVILHRDGEASFDLILMDKEMPERDG						
• • • • •	• • • • •	• • • • •	• • • • •	• • • • •	• • • • •	• • • • •	• • • • •
VXDDPVIRKLHEIIKSIGGISQAKNGEAEVNIHRDGNASFDLILMDKEMPERDG							
10	20	30	40	50			

90	100	110	120	130	140
VSTIKXLREMKGTSMIVGVT	SVADQEEERKAFMEAGLNHCLEKPLTKAKIFPLISHLFDA				
• • • • •	• • • • •	• • • • •	• • • • •	• • • • •	• • • • •
LSATKKLREMKVTSMI	IGVTTLADNEEERKAFMEAGLNHCLAKPLSKAKILPLINNLMDA				
60	70	80	90	100	110

FIG. 10 CONT'D

FIG. 11

GTAATCGGACTCACTATAGGGCACGCGTGGTCGACGCCCGGGCTGTCCTCATTCGTAATTGGGCCCAATGGGCTACT

CACTAGTAGGGCACGGTGGTCG -> ATDZ2F

Spei

AAACACAGTTTCACGATTGTTTTTTTTTTTTTTTAAATTTTAAACATGTATGTGGGATATTGGCTATAAAATTATG 156

TAAAAAATTCACGATAGATTGTTGAATTTTGTGAATTTTCGAGTTAAATAATCTTCAAAATTACCTCACATTACAAAA 234

GGTAGAACTGTTGAAAACTAATGCTCTATATAAACACTAGACAATAACAAAATACGTAATCGTAAAGAACCTAAATT 312

ATGATTTTATTATCTTCTTCCCTTTTCCGTGAGTATAAGCCATTTTTCATAGTAAGCATTACGAATACGACATTG 390

AAACACTACTGACATATATAAAGTAGTAGATTTTGTGATGGTTAACTTGTATGCTTAATTTGCTTAAGCATGAACCTCAATG 468

CTTTTATAAAAGTACTTCATCGAGAATATTCCTCGTTCTATCTACTAGCAGAAGGGTTCGATAGTGATTTTACACCGTTC

AAACAACACCTTTAAACCCCAAAAACCAAGATCTAACTTGATTATACATTTCTGTCTAAATTATCAA 624

ATAACATACCTCTTTTGTCTTACTTATAAACGATATGAAGAAATAAATAAAAGACATAGAAATCTTATTATGATCT 702

AGAAAGAAATTAAGAAATATATATATATTTTTCATTTCTACTCATGTTTCTTATACATTTCTTTAAATTTGTT	780
CACCATTTTGATTTACTTGTCTTCATATAGTTTGTGTTATACAACTCACTTAGAATAATGTAGATTACATTTTCAGCCAA	858
ATTTCATGTAAGATGCTTTTCTTTGTGATGTTTAAATGCTTTCTTTTTCACATTTTCTTTTCTTAACTATAAAT	936
CTTGATCGAATGCCCTTCTTTAGAACATAAGATCTTCTTTTAAATCCAAATCGTAGGCCACTATTTCATTATACT	1013
TATGTAATATATGTGATACAGATACATNTATATACAAATTAACACGAAACCATACATGCACGGTGTGATCACACACG	1092
CACACACATAGAAACATAAACACCGCAATAATTCTTATACAGTTTAAATTTTCACTTTTAACTTTTCTTTTCTTTTGGT	1170
18/27	
NcoI	
<- CTTGAGAGAAAAGAAATCGACCATGGA	
GAAGATTCTTGAGAGAAAAGAAATCGAAGATCGCAACAAAATCCACCGGAGGTACCGAGAAAACCAAGTCGATAGAAG	1248
M A T K S T G G T E K T K S I E V	
TGAAGAAGAACTAATCAACGTGTTGATCGTGTGATGATCCATTAAACCGTAGACTCCACGAGTGTCTCATCAAAA	1324
K K K L I N V L I V D D D P L N R R L H E C H Q ->	

FIG. 11 CONT'D

pDZ2AT3-GUS-SCV

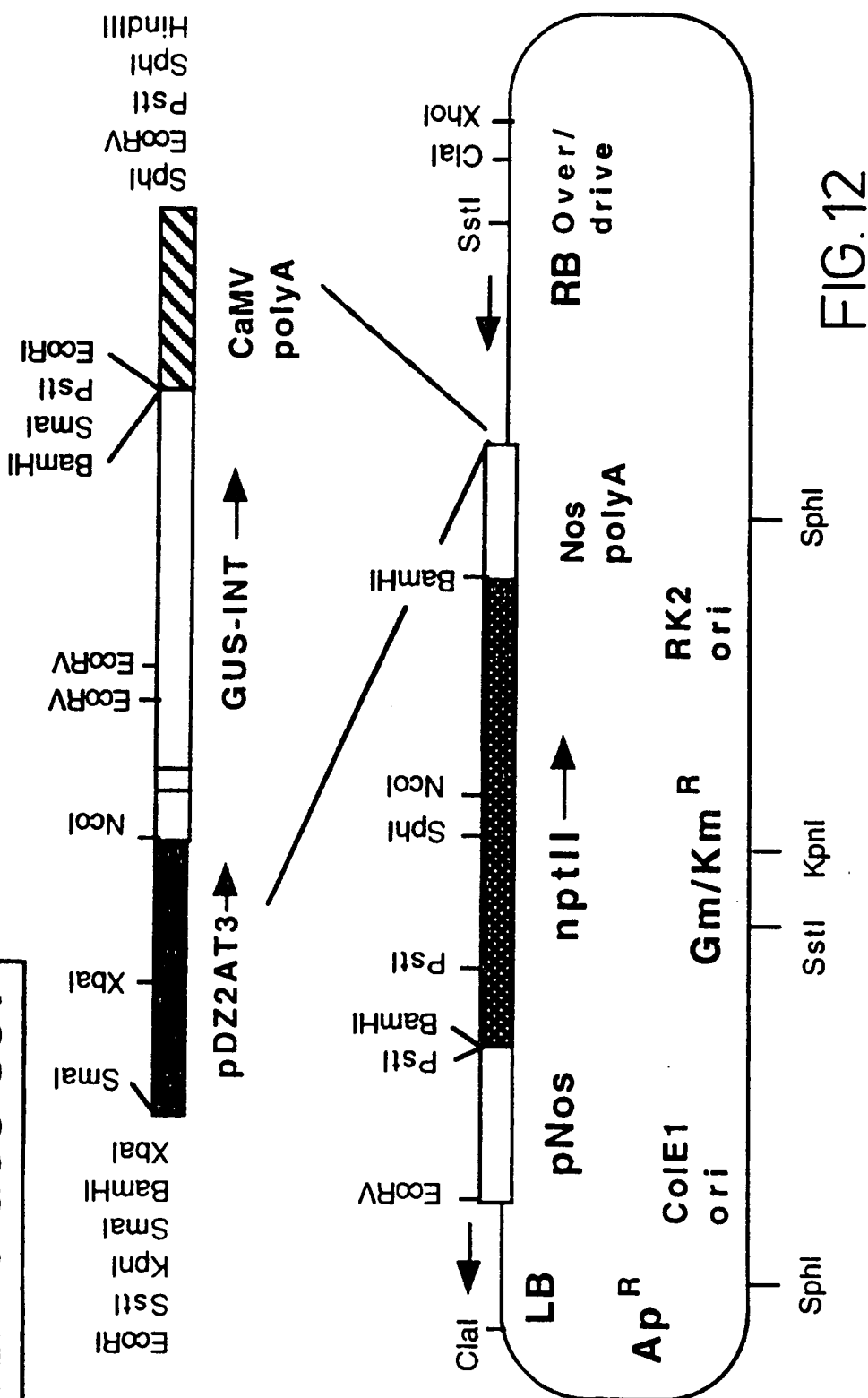


FIG. 12

20/27

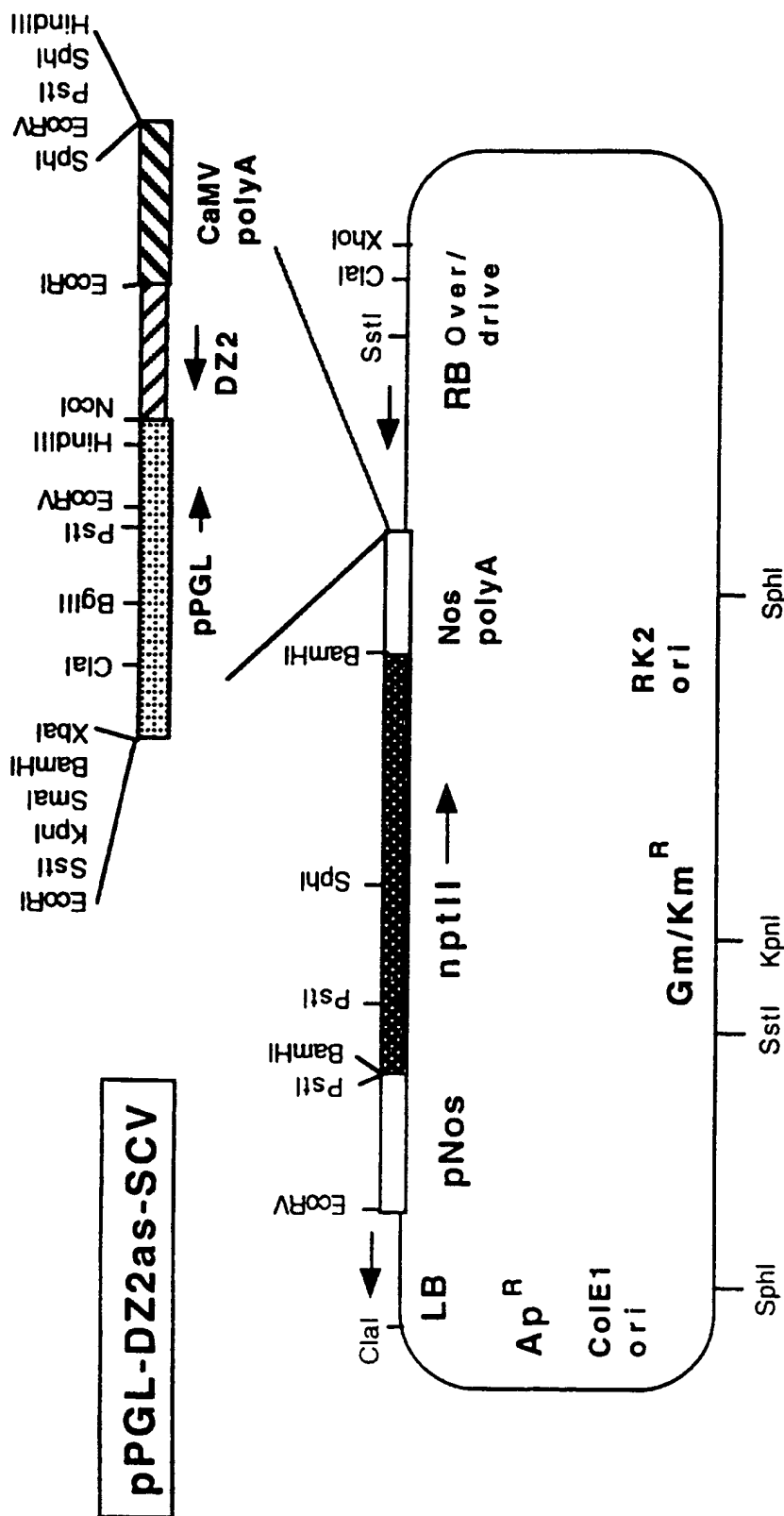


FIG. 13A

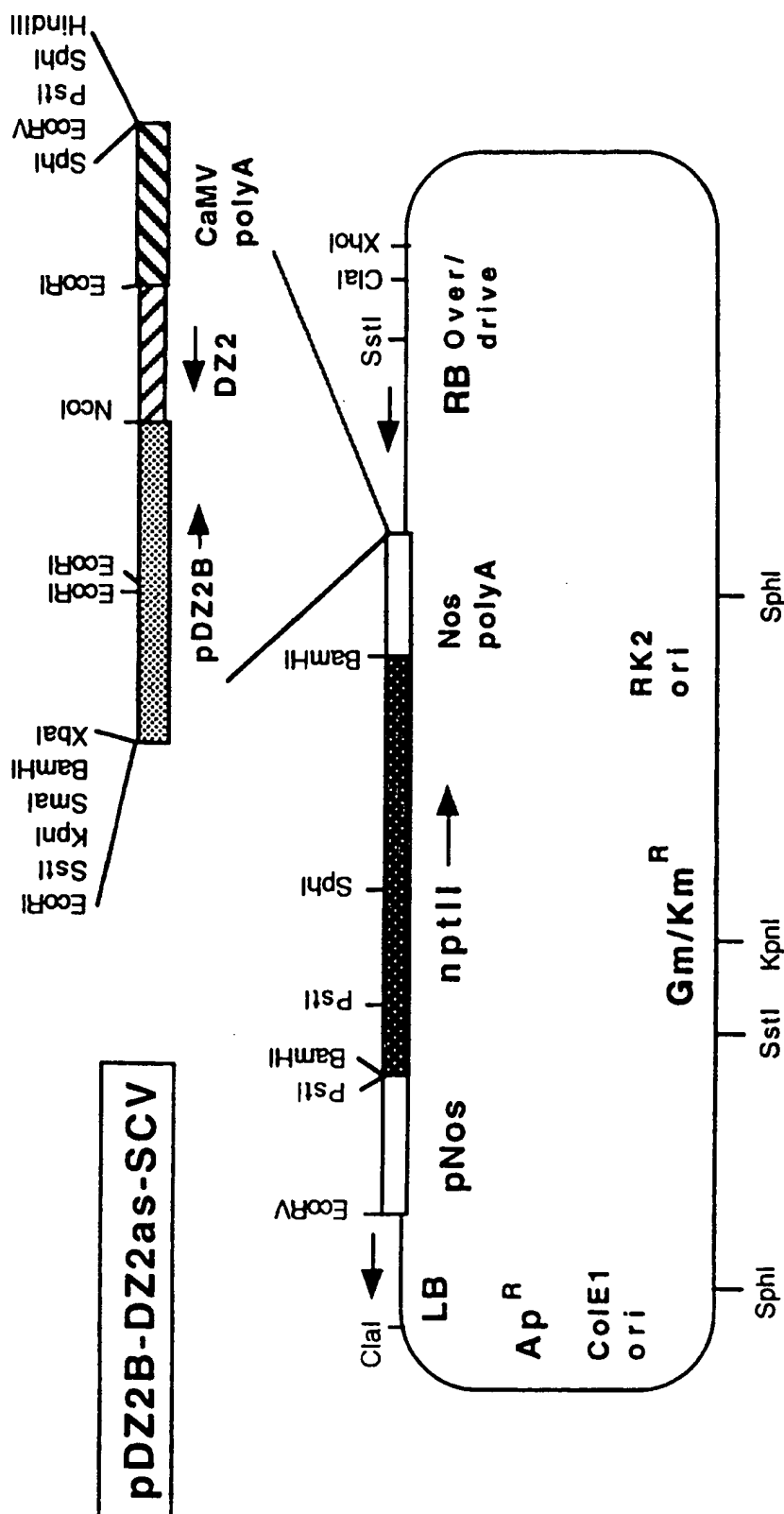


FIG. 13B

22 / 27

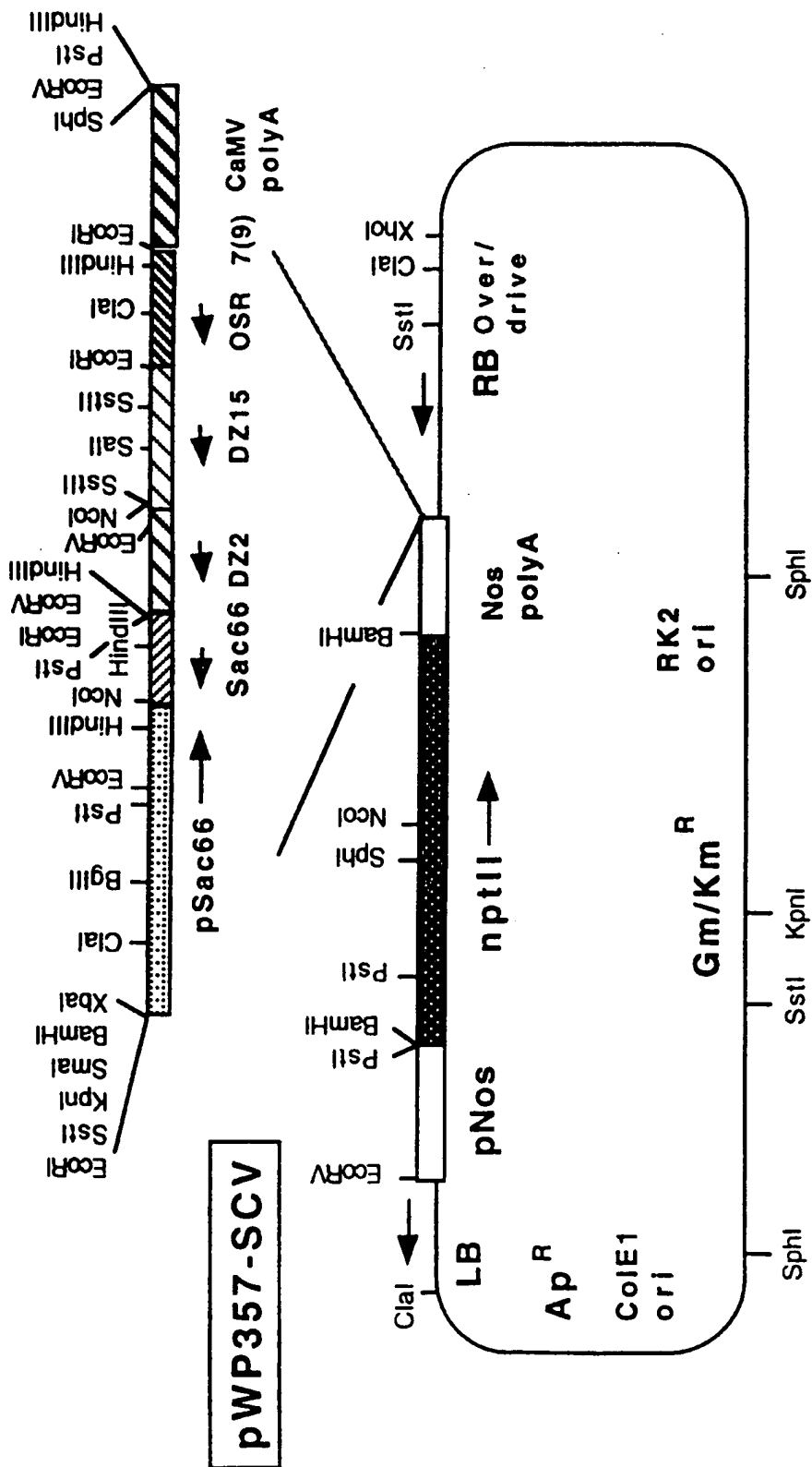


FIG.14

FIG. 15

GGCATCAGGAGGTACCCGTAATAATCCACCATACAACAAGTTCTGTGAAAGTCTCCCA 60

AAACTGCAAGAGTCTCATATTAGTTCTTACTCTCAGAAATAAACACAGTGTCTTGAA 120

AAGATTAGCGTTTCAAAACCCGAAATGGCCCGTTGTTCATGGAAAGTCTTGCTATTTTCTTA 180

M A R C H G S L A I F L 12

TGCGTTCTTTTGATGCTCGCTTGCTGCCAAGCTTTGAGTAGCAACGTAGATGATGGATAT 240

C V L L M L A C C O A L S S N V D D G Y 32

GGTCATGAAGATGGAAGCTTCGAAACCGATAGTTTAAATCAAGCTCAACAACGACGACGAC 300

G H E D G S F E T D S L I K L N N D D D 52

GTTCTTACCTTGAAAAGCTCCGATAGACCCCACTACCGAATCATCAACTGTTAGTGTTCG 360

V L T L K S S D R P T T E S S T V S V S 72

AACTTCGGAGCAAAAGGTGATGGAAAACCGATGATACTCAGGCTTTCAGAAAGCATGG 420

N F G A K G D G K T D D T Q A F K K A W 92

AAGAAGGCATGTTCAACAAATGGAGTGACTACTTTCTTGATTCTTAAGGGAAGACTTAT 480

K K A C S T N G V T T F L I P K G K T Y 112

CTCCTTAAGTCTATTAGATTCAGAGGCCCAATGCAAATCATACGTAGCTTCCAGATCCTA 540

L L K S I R F R G P C K S L R S F Q I L 132

GGCACTTTATCAGCTTCTACAAAACGATCGGATTACAGTAATGACAAGAACCACCTGGCTT 600

G T L S A S T K R S D Y S N D K N H W L 152

24/27

ATTTGGAGGACGTTAATAATCTATCAATCGATGGCGGCTCGCGGGGATGTTGATGCG 660
 I L E D V N N L S I D G G S A G I V D G 172

 ACGGAAAAATCTGGTGCAAACTCATGCAAAATCGACAAATCTAAGCCATGCACAAAA 720
 N G K I W W Q N S C K I D K S K P C T K 192

 GCGCCAACGGCTCTACTCTCTACAACCTAAACAATTGGAATGTGAAGAAATCTGAGAGTG 780
 A P T A L T L Y N L N L N V K N L R V 212

 AGAAATGCACAGCAGATTCAGATTTTCGATTGAGAAATGCAACAGTGTGATGTTAAGAAT 840
 R N A Q Q I Q I S I E K C N S V D V K N 232

 GTTAAGATCACTGCTCCTGGCGGATAGTCCCAACACGGATGGTATTCATATCGTTGCTACT 900
 V K I T A P G D S P N T D G I H I V A T 252

 AAAACATTTCGAATCTCCAATTTCAGACATTGGGACAGGTGATGATTCATATCCATTGAG 960
 K N I R I S N S D I G T G D D C I S I E 272

 GATGGATCGCAAAATGTTCAAATCAATGATTTAACTTGGGCCCCCGTCATGGCATCAGC 1020
 D G S Q N V Q I N D L T C G P G H G I S 292

 ATTGGAAGCTTGGGGATGACAAATTCCAAAGCTTATGTATCGGGAATTAATGTGGATGGT 1080
 I G S L G D D N S K A Y V S G I N V D G 312

FIG. 15 CONT'D

25/27

GCTACGCTCTCTGAGACTGACAATGGAGTAAGAAATCAAGACTTACCAGGGAGGGTCAGGA 1140
A T L S E T D N G V R I K T Y Q G G S G 332

ACTGCTAAGAACAATTAAATTCCAAACATTCGTATGGATAATGTCAAGAAATCCGATCATA 1200
T A K N I K F Q N I R M D N V K N P I I 352

ATCGACCAGAACTACTGCGGACAAGGACAAATGCGGAACAACAAGAAATCTGCGGTTCAAGTG 1260
I D Q N Y C D K D K C E Q Q E S A V Q V 372

AACAATGTCGTCCTATCGGAACATACAAGGTACGAGCGCAACGGATGTGCGGATAATGTTT 1320
N N V V Y R N I Q G T S A T D V A I M F 392

AATTGCAGTGTGAAATATCCATGCCAAGGTATTGTGCTTGAGAAATGTGAACATCAAAGGA 1380
N C S V K Y P C Q G I V L E N V N I K G 412

GGAAAAGCTTCTTGCAAAAATGTCAATGTTAAGGATAAAGGCACCGTTTCTCCTAAATGC 1440
G K A S C K N V N V K D K G T V S P K C 432

CCTTAATTACTAAGTTGATTATGTAATATACATAAAATACGTATTATATGTGGTTATAGAT 1500
P 433

GCCATCTATATCCCTTATCTACGATTGATTCTCGATATATAGAAACTAAGGATTAT 1560

GGGAATATACATAACAATAGTTGAGATAATTGTTGCTCTTGATATGTTCACTGAAGTTGA 1620

TTGCTTGTCACGAATAAATGAATAATGTCATTGTC 1657

FIG. 15CONT'D

26/27

0009456 00 000000 000000

aggtgaccggttgctgatggcaatgtgctgggtcaagcgagaggttagacgggtggcttgaga
V T V A D G N V L V K R E V D G G L E T
cagttaaagtcaaattgccagctgtcattagcgccgacttgcggtcaatgagccgcggt
V K V K L P A V I S A D L R L N E P R Y
acgctactctgcccgaatatcatgaaggccaagaagaagcccatcaaaaagctcacagcca
A T L P N I M K A K K K P I K K L T A T
cagatgtcgggtgtggacttggcgccacgtcaacaagtgttgagcgtagaagacccgcca
D V G V D L A P R Q Q V L S V E D P P T
ccagacaggctggttccattgtgcctgatgtcgacactctcatcaccaagttgaaagaaa
R Q A G S I V P D V D T L I T K L K E K
agggtcattttgtaatgcaatgtcaccaatacagttgttttagttcttacaattcttcgt
G H L *
gaggttttcagctgttaccaataatatTTTTTcaaaatcgattttattttacttgtaatt

taaaagatcaaataattaatacaatgaacatttttgtaacagcaatcttttttttatattt

tggagatttcatcgacttatgtcataattatttttatcaatttattgttggttgtagtg

atataataaagtatgttttctgggtcaaaa

FIG. 16

SECRET

OSR79 306bp

5'
 gg tt gg gt c ga acc at ag gt gg aa ag ct t ct t ct t ct t ct t ct t g a ca aa at c ct t ct gg tt
 L G R T I G G K L L S L S L D K S S G S

 c gg tt t t t c ag t c c c at c ag g ag t t t c t c t at g g t aa ag ct g ag gt t ca aa at g aa a c t t g
 G F O S H Q E F L Y G K A E V O M K L V

 t c c t gg t aa c t c t g t gg aa c ag t c aa ca a c at t c t at c t t aa at c a c g g aa c t a c at
 P G N S A G T V T T F Y L K S P G T T W

 g gg at g ag at c g at t t c g ag t t c t t gg aa a c a t aa ag t gg c c at c c t at a c t c t c c a t a
 D E I D F E F L G N I S G H P Y T L H T

 c t aa t g t t t a c a c g aa gg c t c t gg ag a ca aa g a c ag c ag t t t c a t c t a t g g t t c g a c
 N V Y T R R L W R Q R T A V S S M V R P

 cc g acc 3'
 D

FIG. 17